DOCUMENT RESUME

ED 393 916 TM 024 926

AUTHOR	Kwon, Myoungsook
TITLE	The Use of Planned Comparisons in Analysis of Variance Research.
PUB DATE	Jan 96
NOTE	16p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (New Orleans, LA, January 1996).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Analysis of Variance; *Comparative Analysis; *Educational Research; *Hypothesis Testing; *Research Methodology
IDENTIFIERS	Omnibus Tests; Orthogonal Comparison; *Planned Comparisons; Post Hoc Tests

ABSTRACT

This paper briefly describes unplanned and planned methods of evaluating differences between means and explains orthogonal versus nonorthogonal contrasts to help the researcher understand a framework of planned comparisons. A small heuristic data set is generated to illustrate the superiority of planned comparisons over omnibus analysis of variance (ANOVA) testing and to compare nontrend orthogonal contrasts and nonorthogonal contrasts. The hypothetical data set is used to show that planned comparisons can detect a significant difference between means or a pair of means while a classical test of variance fails to find any statistical significance. Methodologically, planned comparisons have important advantages over unplanned or post hoc tests when the researcher has theoretical ground for focusing on a particular hypothesis or hypotheses. Researchers need to think about the research topic thoughtfully before they choose an omnibus ANOVA test. When they are not certain of the comparisons they wish to make, omnibus ANOVA tests would be applied. (Contains 1 figure. 5 tables, and 16 reference.) (SLD)

Running head: PLANNED COMPARISONS

The Use of Planned Comparisons in Analysis of Variance Research

Myoungsook Kwon

Texas A&M University

Paper presented at the annual meeting of the Southwest Educational Research Association, New Orleans, January, 1996.

The Use of Planned Comparisons in Analysis of Variance Research

Although classical OVA tests (ANOVA, ANCOVA, MANOVA, MANCOVA) remain popular in educational research, some researchers (Benton, 1989; DuRapau, 1988; Keppel, 1982; Kuehne, 1993; Thompson, 1988; Tucker, 1991) have recently suggested that planned comparisons can be superior alternative to unplanned comparisons or classical ANOVA tests followed by post hoc testing . They all maintained two important advantages of using the planned comparisons. One important advantage of planned comparisons is that planned comparisons can have more statistical power against Type II error (not rejecting the null hypothesis when it is false) than unplanned post hoc tests. In other words, it is possible for a specific comparison to be statistically significant when tested by planned comparisons but not significant when tested by unplanned post hoc comparisons (Pedhazur, 1982, pp. 304-305). Another advantage is that planned comparisons may force the researcher to think thoughtfully about research design. To use planned comparisons, a researcher must carefully consider the possible relationships among groups and then decide to which groups ought to be compared.

The present paper briefly describes unplanned and planned methods of evaluating differences between means and explains orthogonal versus nonorthogonal contrasts to help the researcher understand a framework of planned comparisons. A small heuristic data set is generated to illustrate the superiority of planned comparisons over omnibus ANOVA testing and to compare nontrend orthogonal contrasts and nonorthogonal contrasts.

Classical ANOVA Followed by Post Hoc Unplanned Comparisons

Unplanned comparisons (also called *a posteriori* or post hoc) refer to classical omnibus OVA tests followed by post hoc tests. Classical ANOVA methods traditionally test the null

hypothesis of equality among means. If a statistically significant difference is found in the initial ANOVA test involving more than two groups, then the researcher using classical ANOVA methods needs to conduct post hoc or a posteriori tests to locate where the statistically significant results occur within ways having three or more levels. Many "unplanned" or "post-hoc" tests are available, including Scheffe, Tukey, LSD, or Duncan tests, etc. Post hoc comparisons usually involve all possible comparisons of mea .s, even though researchers may be only interested in testing only a few of these comparisons. When a researcher performs multiple t-tests on all possible pairs of means within one study, there is an inflation of experimentwise Type I error rate, i.e., the possibility of making one or more Type I errors for the set of all possible comparisons. As Thompson (1990) notes,

> When several hypotheses are tested within a single study, the experimentwise error rate may not equal the nominal testwise alpha level used to test each of the separate hypotheses. If all hypotheses are perfectly correlated, then and only then will there be no inflation of experimentwise error rate, because in actuality only one hypothesis is really being tested. If the hypotheses are at all uncorrelated, then there will be at least some inflation of the experimentwise error probability. The inflation is at its maximum when the hypotheses are perfectly uncorrelated. (P. 6)

In order to maintain the a priori Type I error rate (α), statistical adjustments (i.e., Bonferroni-type corrections) are incorporated into unplanned tests of comparisons. Even though this procedure controls the experimentwise error rate, it is extremely conservative and may result in no statistically significant comparisons even when the ANOVA test is statistically significant

(Hinkle, Wiersma, & Jurs, 1988). In other words, these adjustments decrease power against Type II errors. Tucker (1991) states that because a given unplanned post hoc test corrects the alpha level for all the possible comparisons for a given study, even comparisons not of interest to the researcher, unplanned tests have less statistical power against Type II error. Due to this trade-off, many researchers encourage the use of alternative analytical methods for comparing group means.

Planned Comparisons

Planned contrasts refer to comparisons of means (simple or complex) that are of the only interest to researchers and the researchers anticipate that these means might be different (Wang, 1993). Planned (also called "*a priori*" or "focused") comparisons, as stated earlier, have been suggested as a valuable alternative to post hoc tests, for two reasons. In addition to increasing power against Type II errors, planned comparisons tend to make the researcher plan the research more thoughtfully. In planned contrasts, a researcher choose hypotheses carefully and then goes directly to the questions (differences between particular groups) that the researcher is interested in, bypassing the omnibus test. Planned comparisons typically involve the weighting of data by sets of contrasts. Various types of planned comparisons are available, including both orthogonal and nonorthogonal comparisons. Also, the contrasts may be nontrend or trend, depending upon the nature of the research design. If a researcher is interested in the patterns in the group means across all the groups, trend contrasts (i.e., linear, quadratic, cubic, etc.) can be performed instead of nontrend contrasts. Discussion on trend contrasts is beyond this paper and more information can be found in Fisher and Yates (1957, pp. 90-100) and Hicks (1973). This paper focuses on orthogonal or nonorthogonal nontrend contrasts.

Orthogonal versus Nonorthogonal Contrasts

Planned comparisons are uncorrelated or orthogonal when each contrast totals zero and when the sum of the cross-products of each pair of contrasts totals to zero. With orthogonal planned comparisons, complete sets of orthogonal contrasts for different hypotheses can be generated. The orthogonal contrasts for three, four, and five treatment groups are presented in Table 1.

INSERT TABLE 1 ABOUT HERE

Each set of orthogonal contrasts can make K-1 contrasts (hypotheses) which always equals to the number of degrees of freedom for a given effect. In three group case, two contrasts (3 - 1 = 2) are possible, testing two hypotheses ($Ho_{(1)}$: $\mu_A - (\mu_B + \mu_C)/2 = 0$, $Ho_{(2)}$: $\mu_B - \mu_c = 0$). Each contrast costs only one degree of freedom. Both planned contrasts and omnibus effects followed by post hoc tests account for the same variance, but do so in different ways. Figure 1 presents three diagrams to illustrate the difference with a three level way.

Some researchers have disagreed on whether to use orthogonal or nonorthogonal contrasts. Some researchers argue that planned comparisons should not necessarily be orthogonal. Hubery and Morris (1988, p. 576) argue that a researcher should ask interesting research questions and not be restricted by orthogonality. The argument is that orthogonal contrasts may not address all possible interest. Furthermore, it is possible for a set of orthogonal contrasts to test extra hypothesis, which are not of interest, to meet the orthogonal constraints. For example, for a three level way presented in Table 1, the researcher is only interested in the first hypothesis . Once the first contrast (2, -1, -1) is established, to be orthogonal the second

contrast (0, 1, -1) must test an additional hypothesis which is not of interest. Winer(1971, p.175) maintains that "whether these comparisons are orthogonal or not makes little or no difference." However, most researchers believe that orthogonal planned comparisons have special appeal. Keppel (1982, p. 147) explains that orthogonal comparisons are uncorrelated. That is, decisions regarding the null hypothesis of one comparison is not influenced by decisions regarding any other orthogonal comparison. Such independence seems desirable in the analysis of relationships among means. The potential difficulty with nonorthogonal comparisons is interpreting the different outcomes because nonorthogonal contrasts provide some overlapping information. Thompson (1990) suggests three advantages of using orthogonal contrasts. First, the exact testwise and experimentwise error rates are both known to us. Second, interpretation tends to be facilitated since equivocal or ambiguous results are less likely. And third, the logic underlying findings can be better generalized to the practice in popular omnibus OVA applications using balanced designs, since classical omnibus tests in such cases are also perfectly uncorrelated.

A Concrete Heuristic Example

A small heuristic data set was generated to illustrate the superiority of planned comparisons over omnibus ANOVA testing. Suppose the research questions are:

1). Is Group 1 different from Group 2, 3, and 4 in the performance score?

2) Are Group 1 and 2 different from Group 3 and 4 in the performance score?

Table 2 presents the example of one-way case with four groups. Three subjects are evenly assigned to each group, thus creating a balanced design. The dependent variable consisted of hypothetical scores on a measure of performance. The one-way ANOVA procedure tests the null hypothesis H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4$.

INSERT TABLE 2 ABOUT HERE

Table 3 presents the results of an one-way ANOVA tests for these data. The procedure indicated no statistically significant differences among the means of the four groups($F=2.4$, $df=3/8$, $p=0.1433$). Therefore the null hypothesis ($\mu1=\mu2=\mu3=\mu4$) is not rejected, indicating no statistically significant difference in the overall F test. Although a post hoc test is not necessary when statistically no significant results are found in one-way ANOVA, for illustrative purposes post hoc analyses were performed. None of the post hoc or unplanned tests detected that any two group means were significantly different at the 0.05 level.

INSERT TABLE 3 ABOUT HERE

However, if a planned comparison tests are performed in the place of the overall ANOVA, a statistically significant difference between means would be discovered. As an example of nontrend orthogonal contrasts, two sets of contrasts (A and B contrasts presented in Table 2) are applied. Each set is made with three contrasts (e.g., A1, A2, A3, ---). The set A is to address the first research question (Ho: $\mu_1 - (\mu_2 + \mu_3 + \mu_4) = 0$). The set B is to address the second (Ho: $(\mu_1 + \mu_2) - (\mu_3 + \mu_4) = 0$). In fact, contrasts A3 and B1 address the very questions the researcher is interested in.

Table 4 presents the results of nontrend orthogonal comparison tests. In Set A, the contrast between G1,2,3 and G4 is found statistically significant at 0.05 level. Also, the contrast between G1,2 and G3, 4 is found significant in Set B at 0.05 level. The total variance (450)

explained by three contrasts in each set equals to the sum of squares explained in OVA test in Table 2.

INSERT TABLE 4 ABOUT HERE

As an example of nonorthogonal contrasts, two contrasts (C1 and C2 presented Table 2) are applied to address the same research questions used in the orthogonal contrasts. Note that the sum of the cross products of the contrast coefficient in the pair is not zero. Only the first contrast testing the hypothesis Ho: $\mu_1 - (\mu_2 + \mu_3 + \mu_4) = 0$ has statistically significant results at 0.05 level, as noted in Table 5. The sum of variance explained by two contrasts (400 + 300 = 700) exceeds the sum of squares explained by ANOVA (450) because these two contrasts provide overlapping information about the explained variance.

INSERT TABLE 5 ABOUT HERE

Summary

This paper used a hypothetical data to show that planned comparisons can detect a significant difference between means or a pair of means while a classical OVA test fails to find any statistical significance. Methodologically, planned comparisons have important advantages over unplanned or post hoc tests when researcher has theoretical grounds for focusing on a particular hypothesis or hypotheses. Therefore, researchers need to think about the research topic thoughtfully before they merely choose an omnibus ANOVA test. When researchers are uncertain of the comparisons they wish to make, omnibus ANOVA tests would be applied. Kerlinger (1986, p. 219) concedes that a posteriori comparisons can be useful in research "for

exploring one's data and for getting leads for future study" but suggests that planned or priori comparisons are "perhaps more important scientifically." To encourage the use of planned comparisons, an in-depth explanation about the methods should be introduced in statistical textbooks.

REFERENCES

Benton, R. (1989, November). Planned comparisons as better alternatives to ANOVA omnibus test. Paper presented at the annual meeting of the Mid-South Educational Research Association, Little Rock, AR. (ERIC Document Reproduction Service No. ED 312 296)

DuRapau, T.M. (1988, November). Benefits of using planned comparisons rather than post hoc tests: a brief review with examples. Paper presented at the annual meeting of the Mid-South Educational Research Association, Louisville, KY. (ERIC Document Reproduction Service No. ED 203 490)

Fisher, R. A., & Yates, F. (1957). Statistical tables for biological, agricultural and medical research (5th ed.). New York: Hafner Publishing Co.

Hicks, C. R. (1973). Fundamental concepts in the design of experiments. New York; Holt, Rinehart and Winston.

Hinkle, D. E., Wiersma, W. & Jurs, S. G. (1988). Applied statistics for the behavioral sciences (2nd ed.). Boston: Houghton Mifflin Company.

Huberty, C. J. & Morris, J. D. (1988). A single contrast test procedure. Educational and Psychological Measurement, 48, 567-578.

Keppel, B. (1982). Design and analysis: A researcher's handbook. Englewood Cliffs, NJ: Prentice Hall.

Kerlinger, G. N. 1986). Foundations of behavioral research (3rd ed.). New York: Holt, Rnehart and Winston.

Kuehne, C. C. (1993, November). The advantages of using planned comparisons over post hoc tests. Paper presented at the annual meeting of the Mid-South Educational Research Association, New Orleans, LA. (ERIC Document Reproduction Service No. ED 364 597)

Pedhazur, E. J. (1982). Multiple regression in behavioral research: Explanation and prediction (2nd ed.). New York: Holt, Rinehart and Winston.

Thompson, B. (1988). The importance of planned or focused comparisons in OVA research. Measurement and Evaluation in counseling and Development, 21, 99-101.

Thompson, B. (1990, April). Planned versus unplanned and orthogonal versus nonorthogonal comparisons: The neo-classical perspective. Paper presented at annual meeting of the American Educational Research Association, Boston. (ERIC Document Reproduction Service No. 318 753).

Thompson, B. (1994). Planned versus unplanned and orthogonal versus nonorthogonal contrasts: The neo-classical perspective. In B. Thompson (Ed.), Advances in social science methodology (Vol. 3, pp. 3-27). Greenwich, CT: JAI Press.

Tucker, M. L. (1991). A compendium of textbook views on planned versus post hoc tests. In B. Thompson (Ed.), Advances in educational research: substantive findings, methodological developments (Vol. 1 pp. 107-118). Greenwich, CT: JAI Press.

Wang, L. (1993). Planned versus unplanned contrasts: Exactly why planned contrasts tend to have more power against Type II error. Paper presented at annual meeting of the Mid-South Educational Research Association, New Orleans, LA (ERIC Document Reproduction Service No. ED 364 598).

Winer, B. J. (1971). Statistical principles in experimental design (2nd ed.). New York: McGraw-Hill.

Figure 1
Variance explained by three methods

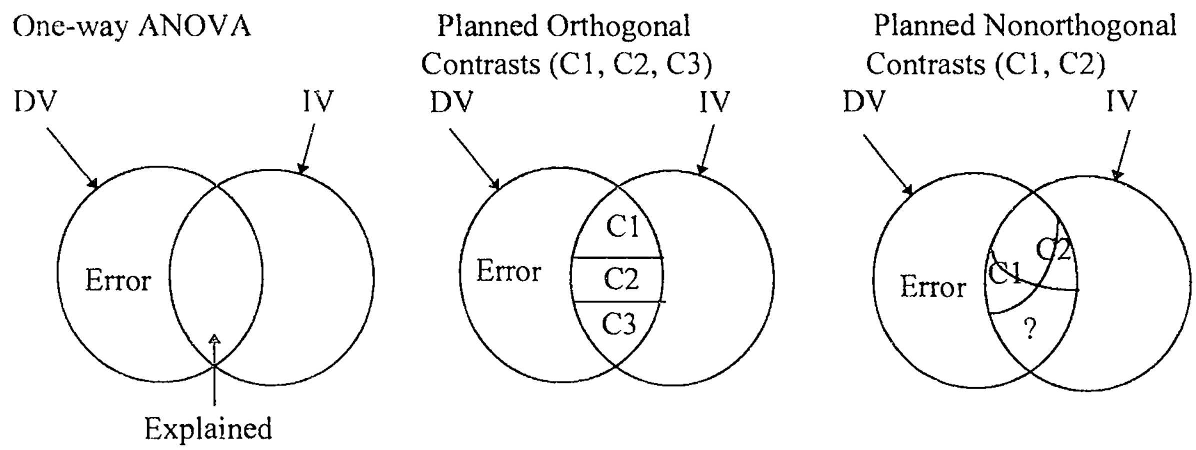

Table 1
Examples of Orthogonal Planned Contrasts

3 Levels

H_A: Is any one group (A in this case) different from other groups?

	A	B	C	
C1	2	-1	-1	(* The hypothesis that a researcher wish to test)
C2	0	1	-1	

4 Levels

H_{A1}: Is any one group (A in this case) different from other groups (B, C, & D)?

H_{A2}: Is there any difference between pairs of treatment groups (A & B vs C, D)?

H_{A1}

	A	B	C	D
C1	3	-1	-1	-1 *
C2	0	2	-1	-1
C3	0	0	1	-1

H_{A2}

	A	B	C	D
C1	-1	-1	1	1 *
C2	0	0	-1	1
C3	-1	-1	1	1

5 Levels

H_{A1}. Is any one group (A in this case) different from other groups (B, C, D, & E)?

H_{A2}: Is any pair of groups (A and B in this case) different from other groups (C, D, & E)?

H_{A1}

	A	B	C	D	E
C1	4	-1	-1	-1	-1 *
C2	0	3	-1	-1	-1
C3	0	0	2	-1	-1
C4	0	0	0	1	-1

H_{A2}

	A	B	C	D	E
C1	-3	-3	2	2	2 *
C2	0	0	0	1	-1
C3	0	0	2	-1	-1
C4	1	-1	0	0	0

Table 2
A Hypothetical Data for Analysis

ID	Group	DV	A1	A2	A3	B1	B2	B3	C1	C2
1	1	10	-1	-1	-1	-1	-1	0	-1	-1
2	1	20	-1	-1	-1	-1	-1	0	-1	-1
3	1	30	-1	-1	-1	-1	-1	0	-1	-1
4	2	10	1	-1	-1	-1	1	0	-1	-1
5	2	20	1	-1	-1	-1	1	0	-1	-1
6	2	30	1	-1	-1	-1	1	0	-1	-1
7	3	20	0	2	-1	1	0	-1	-1	1
8	3	25	0	2	-1	1	0	-1	-1	1
9	3	30	0	2	-1	1	0	-1	-1	1
10	4	30	0	0	3	1	0	1	3	1
11	4	35	0	0	3	1	0	1	3	1
12	4	40	0	0	3	1	0	1	3	1

Table 3
Classical ANOVA Analysis Results

Source	SOS	DF	MS	F	Sig F	Effect Size
Explained	450	3	150.00	2.4	0.1433	47.37 %
Residual	500	8	62.50			
Total	950	11				

Table 4

Results from Planned Orthogonal Contrasts (set A)

Source	SOS	DF	MS	F	Sig F	Effect Size
G1 vs G2 (A1)	0	1	0.00	0.0		0 %
G1,2 vs G3 (A2)	50	1	50.00	0.8		5 %
G1,2,3 vs G4 (A3)	400	1	400.00	6.4	< .05	42 %
Residual	500	8	62.50			
Total	950	11				

Results from Planned Orthogonal Contrasts (set B)

Source	SOS	DF	MS	F	Sig F	Effect Size
G1,2 vs G3,4 (B1)	300	1	300.00	4.8	< .05	31 %
G1 vs G2 (B2)	0	1	0.00	0.0		0 %
G3 vs G4 (B3)	150	1	150.00	2.4		16 %
Residual	500	8	62.50			
Total	950	11				

Table 5

Results from Planned Nonorthogonal Contrasts (set C)

Source	SOS	DF	MS	F	Sig F
G1 vs G2,3,4(C1)	400	1	400.00	6.4	< .05
G1,2 vs G3,4 (C2)	300	1	300.00	4.8	
Residual	500	8	62.50		